DOCUMENT RESUME

ED 262 049 TM 850 508

AUTHOR Wilson, Ann Jarvella
TITLE Knowledge for Teachers: The Origin of the National Teacher Examinations Program.
PUB DATE Apr 85
NOTE 31p.; Paper presented at the Annual Meeting of the American Educational Research Association (69th, Chicago, IL, March 31-April 4, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Historical Materials (060)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Competency Based Teacher Education; *Educational History; Elementary Secondary Education; *Evaluation Criteria; Evaluators; Higher Education; *National Competency Tests; Public Opinion; Standardized Tests; Teacher Certification; Teacher Effectiveness; *Teacher Evaluation; *Teacher Qualifications; Teacher Stereotypes; *Test Construction; Testing Programs
IDENTIFIERS *National Teacher Examinations

ABSTRACT

The National Teacher Examinations (NTE), were first administered in 1940. Teacher examinations had existed in the United States since colonial New England. During the late 1920's, intelligence tests became popular; teaching aptitude, subject matter competence, and knowledge of testing were also measured. In 1931, a cooperative program of the Teachers College Personnel Association and the American Council on Education tested entering education majors. Quality was found to vary considerably. A large-scale evaluation program, initiated by the Carnegie Foundation for the Advancement of Teaching, tested high school and college students in Pennsylvania. The Cooperative Test Service (CTS) of the American Council on Education, organized in 1930, provided teacher tests for participating school districts. The tests covered English comprehension and expression, reasoning, general culture, professional information, contemporary affairs, and relevant subject areas. The director, Ben Wood, promoted use of the test as one criterion for teacher selection, while making statements about the mediocrity of teachers. The CTS's first NTE, developed in 1939,resembled its earlier versions. It stressed basic intellectual and communication skills, cultural and contemporary background, and professional information. These areas are still important in the current NTE, which is administered by Educational Testing Service. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED262049

KNOWLEDGE FOR TEACHERS:

THE ORIGIN OF THE NATIONAL TEACHER EXAMINATIONS PROGRAM

Ann Jarvella Wilson

Department of Education
Carroll College
100 North East Avenue
Waukesha, Wisconsin 53186

Paper presented at the Annual Meeting
of the American Educational Research Association
Chicago, April 1985

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Wilson, Ann J.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

DRAFT COPY:

Do not quote or copy without permission.
Comments are welcome.

TM 850 508

Knowledge for Teachers:

The Origin of the National Teacher Examinations Program

ANN E. JARVELLA WILSON, Carroll College, Waukesha, Wisconsin

This paper explores those factors which in 1940 culminated in the original NTE project and influenced the content and form of the first exams. Based on the actual examinations and other primary documents prepared for candidates, users, private foundations and other sponsors, the study describes and interprets test content within the social context of its development. The construction, scope, and substance of those first exams reflect assumptions about assessment, knowledge, and teachers which were held by test developers of that period and which have influenced the direction of standardized teacher testing for the past fifty years.

Introduction

Heightened public and professional concern about the quality of our nation's schools and the competence of the teachers who work within them has led to growing interest in using tests for teacher assessment and/or certification. Although agencies in some areas have constructed their own instruments, the use of the National Teacher Examinations has become increasingly popular. The growth in test use has been accompanied by some controversy and debate, but for the most part both the history of the program and the content of the exams have received little attention

Currently prepared by the Educational Testing Service, this battery of standardized tests for prospective teachers was first officially administered in March 1940. This paper, which is drawn from a larger analytic history of the NTE program,[1] explores the social context and the specific content of those original exams[2] designed to assess knowledge common to all teachers. Initiated to assist urban school superintendents in selecting candidates from an oversupply of teachers believed to vary considerably in training and ability, the original tests assessed those aspects of general knowledge, professional information, and intellectual and basic skills which the administrators and first test developers believed all teachers should possess. The construction, scope, and substance of those first examinations reflect assumptions about the competence, assessment, and knowledge of teachers which were held by test developers of that time and which appear to have guided the direction of teacher testing over the past fifty years.

[1] Ann Jarvella Wilson, "Knowledge for Teachers: The National Teacher Examinations Program, 1940 to 1970," unpublished Ph.D. dissertation, University of Wisconsin, 1984, available from University Microfilms International as No. 84-14265.

[2] Access to the NTE history files was provided by Educational Testing Service, and permission to cite excerpts from those files was granted by ETS and the American Council on Education--sponsor of the NTE program from 1940 to 1949. The opinions and conclusions expressed in this paper are those of the author, and neither ETS nor ACE has participated in nor bears any responsibility for this study.

Early Certification by Examination

Originally developed to license or certify teachers for medieval European church-sponsored universities, examinations for teachers were first used in the United States in colonial New England. Potential school masters needed to convince their local ministers of the soundness of their faith, of their moral as well as scholastic accomplishments. With the expansion of public schooling, a reliance on teacher examinations spread throughout the nation.

Although the secular certifying agencies varied somewhat from locale to locale, most of the early exams continued to assess teaching candidates in terms of their moral character and their ability to teach the common school subjects. By the late 1800's, a few examination boards attempted to assess professional or pedagogical knowledge as well.

With the expansion of college and normal school training programs, certification by examination eventually was supplanted by the acceptance of credentials showing the completion of a prescribed sequence of school or college courses. Although mandatory county and state-wide examining declined over the first few decades of the twentieth century, teacher testing persisted in a number of large cities. Because of the demand for urban teaching positions, hiring authorities often required examinations in addition to the completion of a training program. In rural areas, tests continued to be used to grant "emergency" low-grade licensure to candidates who had not completed professional training.

Although there were some outspoken critics of the quality of the early examinations, test use was usually seen as unavoidable. Despite their expansion, the training institutions did not graduate sufficient candidates to fill the burgeoning number of positions. The use of the exams was seen as a practical way to achieve some common, albeit minimal, standard of teacher competence.

Standards and Testing within Teacher Training Institutions

During the late 1920's, concerns about the quality of teachers and of teacher education intensified. A lowered national birthrate had begun to produce a leveling off in the school population. The relatively higher salaries and improved conditions instituted to combat severe teacher shortages during the war attracted increasing numbers of qualified novices as well as some former teachers who had left earlier for better paying positions in government and industry.

With the depressed economy of the early 1930's, non-academic job opportunities were scarce, and the number of teachers and teacher candidates who left for other occupations was greatly reduced. Thus, from a combination of factors, there was actually a surplus of certified teachers in many areas of the country, particularly in the larger cities. Some teacher educators used the oversupply as a rationale to call for the elimination of certification by examination. Others recommended the imposition of higher standards in teacher training. Some called for increased institutional testing.

At about this same time, nationwide emphases on efficiency and accountability and the thriving intelligence and achievement testing movements were exerting considerable pressure upon the nation's schools. Teacher education was not spared. Recommendations were made for trainee assessment in such qualities as "functional competency in teaching areas, general education, intelligence, professional aptitude and attitude, scholarship, and physical and mental health."[3] Despite a number of studies which indicated little relationship between intelligence test scores and other "measures" of success in either the training program or teaching itself, the administration of group IQ tests and the selection of teacher education candidates by intelligence test score

[3] Alexander J. Stoddard, "A Hundred Thousand New Teachers Every Year," in American Association of Teachers Colleges, Seventeenth Yearbook (Menasha, WI: Banta, 1938), p. 88.

became quite popular. Instruments were also developed to measure trainees' teaching aptitude, subject matter competence, and knowledge of testing.

Standardized Testing Programs for Teaching Candidates

Most of these tests had limited use outside of their authors' schools, but a few received more widespread attention. Three large scale teacher testing projects appeared in the late 1920's and early 1930's--those initiated by the Bureau of Public Personnel Association, by the Teachers College Personnel Association, and by the Carnegie Foundation for the Advancement for Teaching--and were antecedents of the National Teacher Examinations program.

In the late 1920's, the Bureau of Public Personnel Administration offered for sale nationally several tests for elementary and high school teachers. Originated in 1922 for his doctoral research at Teachers College, Columbia University, the tests were modified by F.B. Knight and others at the University of Iowa.[4] The exams for elementary school teachers contained teaching aptitude sections on professional judgment, theory and practice of teaching, reading comprehension, social information, school and class management, and professional information. Also included were subject placement exams which concentrated upon the pedagogical issues involved in teaching reading, arithmetic, spelling, and writing and which were "devised and standardized on the assumption that knowledge of techniques of teaching a given subject distinguishes a good teacher from a poor teacher in that subject."[5] Although most of the test development and standardization involved students in the teachers colleges, subsequent use and test sales were aimed at school system personnel. School superintendents were

[4] For example, see F.B. Knight et al., "Standardized Tests for Elementary Teachers," Public Personnel Studies 4 (October 1926), pp. 279-298; F.B. Knight, "The Selection of High School Teachers," in National Society of College Teachers of Education, Fifteenth Yearbook (Chicago: University of Chicago Press, 1926), pp. 35-37.

[5] "Information and Data Regarding Bureau Tests Previously Published: Standardized Tests for Teachers," Public Personnel Studies 5 (September 1927), p. 196.

urged to use the tests to evaluate teachers already employed as well as those applying for new positions.

In 1931, a cooperative testing program was initiated by the Teachers College Personnel Association, an affiliate of the American Association of Teachers Colleges and of the American Council on Education.[6] Growing out of an "experiment" in teachers college administration begun in the mid-1920's at Colorado State Teachers College, it made intelligence tests, achievement exams, and personnel record materials available for sale to teachers colleges. Purchasers were encouraged to return their test results so these could be analyzed and included in cooperative reports.

Used in several ways by the participating colleges--some one hundred six between 1931 and 1939--the tests' primary function was to evaluate the quality of entering students.[7] The exams revealed that this quality varied considerably from individual to individual and from college to college. Although testing teacher candidates following completion of professional training was not suggested specifically in any of the Teachers College Personnel Association literature, the results of the cooperative testing program were used later by those who advocated testing teachers prior to employment.[8]

The third large scale evaluation program--that initiated in Pennsylvania by the Carnegie Foundation for the Advancement of Teaching--was not originally conceived to assess teaching candidates. However, it resulted in extensive teacher testing and led directly to the National Teacher Examinations program.

[6] J.D. Heilman, "The Cooperative Program of the Teachers College Personnel Association," in American Association of Teachers Colleges, Eighteenth Yearbook (Menasha, WI: Banta, 1939), pp. 71-79.

[7] Nora A. Congdon, The 1940 Report on the Cooperative Testing Program of the Teachers College Personnel Association (Greeley: Colorado State College of Education, 1940), p. 8.

[8] Ben Wood, later the director of the NTE program, cited the cooperative testing results along with his own research in "Teacher Selection: Tested Intelligence and Achievement of Teachers-in-Training," Educational Record 17 (July 1936), pp. 374-387.

The Pennsylvania Study

Begun in 1925 as a Carnegie Foundation project funded by the Carnegie Corporation,[9] the Study of the Relations of Secondary and Higher Education in Pennsylvania evolved from concerns of Pennsylvania's college and state level school administrators as well as from an earlier Carnegie study comparing United States and European educational institutions. In an attempt to assess the quality of the Pennsylvania secondary and higher educational systems, lengthy batteries of objective tests were administered to selected high school and college students between 1928 and 1934.

Directing the study along with William Learned of the Carnegie Foundation staff was Ben Wood, a national authority on objective testing and the director of Collegiate Educational Research at Columbia University. His doctoral dissertation, done as a student of Edward Thorndike and published with an introduction written by Lewis Terman, involved the construction and analysis of a variety of objective college-level content examinations.[10] He had developed standardized exams for Columbia in both law and medicine and, with a previous grant from the Carnegie Corporation, had produced high school level objective examinations for New York State in physics and several foreign languages.

In May 1928, all graduating high school seniors in Pennsylvania (approximately 27,000) were administered a massive battery of commercial intelligence and achievement tests. Special comprehensive exams developed by Wood and his colleagues at Columbia to "test memory, judgment, and reasoning ability through simple recognition" were given to

[9] Howard J. Savage, "Educational Grants of the Carnegie Corporation and the Carnegie Foundation" in Carnegie Foundation for the Advancement of Teaching, Thirtieth Annual Report (New York: Updike, 1935), p. 92.

[10] Ben D. Wood, Measurement in Higher Education (New York: Teachers College, Columbia University Press, 1923).

approximately 4400 college seniors that same year.[11] Following revision and shortening (down from twelve hours to eight), these exams were administered twice to members of the 1928 high school class who went on to college--first as sophomores in 1930 and then again as seniors in 1932. These same tests were also given to several groups of high school seniors in 1933 and 1934. The exams contained matching, true-false, and multiple-choice items and were designed to assess intelligence, English (spelling, grammar, punctuation, vocabulary, and literature), mathematics, and "general culture" (general science, foreign literature, fine arts, history and social studies).

The tests were intended to measure "knowledge that because of some congenial use or welcome has been retained, worked over, and made effective."[12] Speaking for himself and Wood in a 1933 report to the Carnegie Foundation, Learned stated: "In our judgment, the acquisition, retention, and sound use of thoroughly criticized and related knowledge should be made the foundation of all formal educational effort, just as smoothly functioning knowledge is the obvious secret of every effective educational result."[13]

The major finding of the Pennsylvania study was the "discovery" of great variability in tested knowledge, variability which was exhibited among individuals and among institutions as well as within classes and departments in the same institution. Neither class placement nor school grades necessarily corresponded to knowledge displayed on the tests. For those students who were followed longitudinally and tested repeatedly, time in school did not always correlate with test score gains. In some

[11] William S. Learned and Ben D. Wood, The Student and His Knowledge: A Report to the Carnegie Foundation on the Results of the High School and College Examinations of 1928, 1930, and 1932, Carnegie Foundation for the Advancement of Teaching Bulletin No. 29 (New York: Updike, 1938), p. 372.

[12] William S. Learned, "Testing College Students," in Carnegie Foundation for the Advancement of Teaching, Twenty-Fifth Annual Report of the President and of the Treasurer (New York: Updike, 1930), p. 52.

[13] William S. Learned, "Knowledge as a Factor in Education: The Tests and Their Implications," in Carnegie Foundation for the Advancement of Teaching, Twenty-Eighth Annual Report of the President and of the Treasurer (New York: Updike, 1933), p. 44.

cases, high school seniors showed "more" knowledge than college sophomores or even seniors. Eventually the findings and interpretations of the Pennsylvania study led to extensive nationwide college and graduate testing projects, one of which evolved into the Graduate Record Examination program.

Though not an original focus of the research, the results of the Pennsylvania study became widely used later to decry the academic quality of teachers and teacher candidates. As part of the data analysis, test score comparisons were made of college students in various pre-professional and occupational programs. The prospective teachers about to graduate from college as seniors in 1928 and in 1932 tested particularly poorly. "In both [years] the teachers' average was below the average total score for the entire group and was below all other group averages except those of the business, art, agriculture, and secretarial candidates."[14]

In addition, the scores of the teacher candidates were compared with those of their prospective pupils--seniors in the Pennsylvania high schools. Again there was considerable variability, but the scores of some twelve percent of the high school seniors did exceed the median of the teacher group.[15] From this Learned and Wood concluded that "high-school teaching attracts college students who differ widely in the fundamental quality of their abilities and who fall below a knowledge minimum in a large proportion of cases."[16] Comparison of intelligence scores confirmed the "conclusion as to the limited mental ability of the individuals who are being specially prepared for teaching positions." The teaching candidates tested were seen as "narrower people" with "uninformed and incompetent minds," persons who might be

[14] Learned and Wood, pp. 38-39.

[15] William S. Learned,"Tested Achievement of Prospective Teachers in Pennsylvania," in Carnegie Foundation for the Advancement of Teaching, Thirty-First Annual Report (New York: Updike, 1936), p. 45.

[16] Learned and Wood, p. 43.

suitable elementary teachers but certainly would be inadequate in the high schools.[17]

Most commentaries upon the Pennsylvania study concentrated upon test score variability. Although some reviewers criticized prevailing educational practices, most did not comment, except briefly, upon the teacher scores. Of course, brief mention could be pernicious as in Lewis Terman's classification of some teachers as "congenital ninth graders."[18]

In his lengthy and critical review of the Pennsylvania study, Goodwin Watson[19] of Teachers College, Columbia University, considered the content of the exams as well as the mediocre performance of prospective teachers on them. Although believing the exams were carefully and technically well constructed, he questioned whether they tested the "real" objectives of education. It was "possible," he suggested, "to carry the emphasis on subject matter too far" for other studies had shown "very little relationship between academic scholarship and teaching success."[20] Conceding that potential teachers needed a stronger academic foundation in some areas, Watson recommended raising teacher salaries, adjusting teacher education curriculum, and changing the public's vision of teachers. Higher quality persons would be attracted to teaching, he felt, when teachers were recognized as "leaders in the creation, interpretation, and evaluation of . . . social policies." For the present, poor scores by teacher candidates were inevitable for "only mediocre minds will be interested in classroom routines, tests, marks, and the details of education."[21]

[17] Ibid., pp. 351-353.

[18] Lewis Terman, "An Important Contribution," Journal of Higher Education 10 (February 1939), p. 112.

[19] Goodwin Watson, How Good Are Our Colleges?, Public Affairs Pamphlet, No. 26 (New York: Public Affairs Committee, 1938).

[20] Ibid., p. 29.

[21] Ibid., pp. 29-30. F.B. Knight, author of the Bureau of Public Personnel Administration's teacher tests, also had thought that "too much intelligence" might be disadvantageous "in the rather humdrum and steady grind of elementary school teaching." ["Qualities Related to Success in Elementary School Teaching," Journal of Educational

Learned and Wood claimed that the "eventual solution to the problem of selecting teachers who both [knew] what to teach and [were] themselves educated" would involve better programs and higher standards in the preparatory schools and colleges.[22] However, in their major report on the Pennsylvania study, they put much emphasis upon the use of examinations for teaching candidates. In particular, they advocated that, prior to teacher employment, school personnel require tests such as those then used by the Providence, Rhode Island, school system.[23] What they did not specify was that these examinations had been developed by the Cooperative Test Service under Wood's direction and were based upon the very tests used in the Pennsylvania study.

The Cooperative Test Service

The Cooperative Test Service of the American Council on Education had been officially organized in October 1930 with the appointment of Ben Wood as director. Funded with a ten year grant from the General Education Board, a foundation established with monies donated by John D. Rockefeller, the project was the culmination of almost a decade's involvement in vocational and educational testing by the American Council on Education and several private foundations. Established as a test producing bureau, the service was expected to prepare multiple comparable forms of tests "in each of the major academic subject-matter fields at the senior high school and junior college levels."[24] A number of exams were donated, including those prepared for the Pennsylvania study.[25] Director Wood and other central staff were housed on the Columbia University campus in New York City, but test construction was distributed

Research 5 (March 1922), p. 214.]

[22] Learned and Wood, p. 65.

[23] Ibid., pp. 64-65.

[24] Ben D. Wood, "Ten Years of the Cooperative Test Service," Educational Record 21 (July 1940), p. 371.

[25] Max McConn, "The Co-operative Test Service," Journal of Higher Education 2 (May 1931), pp. 231-232.

around the country, with subject-matter specialties concentrated in various university affiliated testing bureaus.

Although its initial mission was test construction, from the beginning the Cooperative Test Service did more than produce academic examinations. One of its earliest projects, begun in 1932, extended the work of the Pennsylvania study with the development and coordination of the college sophomore testing program.[26] Using certain of the tests donated by the Carnegie Foundation, the first year's program called for the administration of intelligence, English, and general culture exams (general science, foreign literature, fine arts, and history and social studies) to more than 18,000 students in one hundred forty participating institutions.[27]

In that same year, "the school authorities of Providence, Rhode Island, asked the Cooperative Test Service for a special edition of its tests for use as one phase of their teacher selection procedure."[28] The bureau complied, and soon several other cities also began using the exams. By the late 1930's, the test service was providing new versions for teacher selection yearly to some fifteen or twenty cities including Philadelphia, Pittsburgh, and Cleveland.[29]

In 1938, the Cooperative Test Service informed participating school officials that its subsidizing grant would expire in 1940 and thus it could no longer supply them with "new and unused tests suitable for competitive exam purposes."[30] The school superintendents then conferred with the American Council on Education, and a National

[26] Ibid.

[27] J.B. Johnston et al., "The 1932 College Sophomore Testing Program," Educational Record 13 (October 1932), pp. 294-296.

[28] Matthew J. Downey, Ben D. Wood: Educational Reformer (Princeton: Educational Testing Service, 1965), p. 60.

[29] Howard H. Long et al., Principles and Procedures in Teacher Selection (Philadelphia: American Association of Examiners and Administrators of Educational Personnel, 1951), p. 15.

[30] Wood, "Ten Years of the Cooperative Test Service," pp. 374-375.

Committee on Teacher Examinations was appointed.

The National Committee on Teacher Examinations

Composed primarily of superintendents whose urban school systems had been using the tests provided previously, the committee held its first meeting in September 1939. Its ten members were to serve as advisors to a teacher examination project affiliated with the Cooperative Test Service. Superintendent Alexander J. Stoddard of Philadelphia was chosen as chairman. Having directed Providence's schools as superintendent from 1929 to 1937, he had been instrumental in initiating teacher testing there.

Early in the meeting, the history of the project and current accomplishments were reviewed. Funding for the project had been secured by the American Council on Education with a grant from the Carnegie Corporation through the Carnegie Foundation. Under the supervision of Ben Wood as project director, the tasks of constructing, administering, and scoring the exams were assigned to the Cooperative Test Service. Suggestions for exam design and content had been solicited at preliminary conferences for school superintendents, during personal visits to "school administrators and other interested individuals" in seventeen cities nationwide,[31] and from a questionnaire circulated to a larger number of superintendents in the spring of 1939.[32] From these, the Cooperative Test Service staff had developed "expanded outlines" which were shared with the committee at the meeting. The "central core" for all teachers would consist of five parts--English comprehension and expression, reasoning, general culture (history and social studies, current social problems, literature, fine arts, science, and mathematics), professional information, and contemporary affairs. Also

[31] Donald J. Shank, "Minutes of the Meeting of the National Committee on Teacher Examinations, September 23-24, 1939, New York City," confidential unpublished document from NTE History File, p. 3.

[32] Ben D. Wood, "Examinations for Teachers of Business Education," in National Association of Business Teacher-Training Institutions, Proceedings of the Fourteenth Annual Conference, Bulletin No. 22 (Chicago: the Association, May 1941), p. 5.

comprehensive subject area tests would be developed in the traditional high school subjects and for elementary school teachers.[33]

After some consideration, the Committee agreed: (1) that the examinations "should be based primarily upon the objectives of school systems," using descriptions of what "good teachers should know" obtained from "all cooperating school systems;" (2) that the five basic examinations were more important to the administrators than the subject-matter exams in specific fields; (3) that the basic tests should be "functional in nature" and should not overlap in subject-matter with the survey exams; and (4) that included in the basic examinations should be measures of "how much teachers know about the personal, psychological, emotional, and physical development of boys and girls."[34]

Promotion of the Teacher Examination Service

Just two months after the meeting, in November 1939, "an announcement of a teacher examination service"[35] written by Ben Wood was issued by the National Committee. Containing a rationale and justification for the program as well as descriptions of the service and of the exams, it set forth the basic assumptions and standard arguments repeatedly used later by Wood and other exam advocates. Emphasizing the importance of teaching and thus of teacher selection, it pointed to difficulties stemming from a variety of standards at institutions of teacher preparation and from the complexity of factors contributing to good teaching. Great stress was placed upon using the tests as only "one phase" of teacher selection. It was suggested that use of the exams could result in school authorities paying greater attention to other crucial factors because participation in the national program would save the time and expense of constructing,

[33] Shank, pp. 2-3.

[34] Ibid., p. 5.

[35] Ben D. Wood, An Announcement of a Teacher Examination Service (New York: National Committee on Teacher Examinations of the American Council on Education, November 1939).

administering, and scoring local tests.[36] Although the service was admittedly developed primarily to assist superintendents, it was also suggested that "the opportunity to 'register' talents on a national scale" would be advantageous to candidates and to institutions preparing teachers.[37]

Since the mid 1930's, Ben Wood had been attempting to garner teacher educators' support for the testing project. His initial report on the mediocre performance of teaching candidates in Pennsylvania had been originally presented to an audience of teacher trainers. It had described the examination program in Providence and suggested that teacher educators work cooperatively with the Teachers College Personnel Association and the American Council on Education. However, he also had stated that many teaching candidates belonged to "the horde of semi-literates who flaunt their diplomas before the credulous eyes of employer superintendents."[38]

Similarly uncomplimentary and revealing comments by Wood about teachers had been published earlier that year. An article written with F.S. Beers of the University of Georgia[39] had suggested that "large proportions" of teachers could not think and that many could not "even learn or understand . . . bare facts and simple relations,"[40] that children could better learn "if their teachers would condescend both to learn and to teach knowledge appropriate to the abilities, interests, and real needs of their students,"[41] and that ". . . education classes are as much if not more amply populated

[36] Ibid., p. 12.

[37] Ibid., p. 4.

[38] Wood, "Teacher Selection," p. 381.

[39] Ben D. Wood and F.S. Beers, "Knowledge versus Thinking?," Teachers College Record 37 (March 1936), pp. 487-499.

[40] Ibid., p. 487.

[41] Ibid., p. 496.

with morons than other departments . . ."[42] Given published statements like these, it is not surprising that educators viewed Wood and his suggestions with some reservation.

Apparently the Carnegie Foundation was not unaware of the potential animosity with which teachers and their educators might react to Wood's testing proposal. Although he had been attempting to secure funding for a nationwide program for a number of years, the Foundation "anticipated a furious reaction from many teachers and teachers colleges and was not eager to get involved."[43]

The promotional campaign launched for the first test administration was aimed both at attracting support from city school system officials and at reassuring teachers and training personnel. In a presentation to the prestigious yet potentially hostile American Association of Teachers Colleges, Wood delivered an impassioned plea for exam use as "one phase" of teacher selection.[44] "In a world," he argued, "in which peoples squander billions on the gleam in some politician's eye, in which great nations revert to barbarism in worship of a maniacal super-salesman, in which American planes and bombs rain death on the innocent men, women and children of friendly nations, and in which we are confused by dozens of similar paradoxes, the only possible hope for our children lies in having them educated, so far as possible, by persons who are themselves educated. I believe that the wise and judicious use of examinations such as those provided by the National Committee on Teacher Examinations will help assure this boon for our school children."[45]

[42] Ibid., p. 498.

[43] Downey, Ben D. Wood, p. 60. According to Downey, Wood's ETS biographer, he finally did succeed in 1938 in a chance meeting with the president of the Carnegie Foundation, Walter Jessup, in Paris.

[44] Ben D. Wood, "Making Use of the Objective Examination as a Phase of Teacher Selection," in American Association of Teachers Colleges, Nineteenth Yearbook (Menasha, WI: Banta, 1940), pp. 15-19.

[45] Ibid., p. 19.

He stressed that "objective examinations do not and cannot measure the total subtle complex which we call teaching ability"[46] and argued against "the naive error of judging . . . validity . . . in terms of [the tests'] correlation with available criteria of teaching success."[47]

Over the next few years, numerous papers focusing on the examination program were presented at national conferences and published in educational journals and yearbooks. The exams were both praised and condemned. Throughout the debate, those responsible for the tests emphasized a critical relationship between test validity and test content. Again and again, Wood argued that the value of the exams could not be judged by correlating them with "available criteria." Since teaching ability was a complex combination of numerous interacting factors, it was not "reasonable to expect any one of the essential factors to correlate highly with the total complex."[48] Instead, Wood declared that the tests should be judged by how accurately they measured those parts of teaching they were "designed to measure . . ."[49] Although Wood maintained that "the relevance of [the content] to teaching fitness must be left to the judgment of the selecting authorities, acting at least partly on a priori and common sense considerations,"[50] he emphasized the importance of exam construction and content selection.

[46] Ibid., p. 15.

[47] Ibid., p. 16.

[48] Ben Wood, "Dr. Wood's Statement," Progressive Education 17 (March 1940), p. 156.

[49] Wood, "Making Use of the Objective Examination," p. 19.

[50] Ibid.

Construction of the 1940 National Teacher Examinations

The first forms of the National Teacher Examinations were prepared under the direction of the staff of the Cooperative Test Service during the fall and winter of 1939. Following procedures originally used by Wood in the Pennsylvania study, staff editors developed preliminary test outlines and tentative item specifications. The general suggestions gathered earlier from school administrators and from the National Committee on Teacher Examinations were supplemented with data gleaned in analyses of courses of study, textbooks, journal articles, and reports of professional organizations. Outlines were sent to teacher trainers and school system personnel for review and criticism. Incorporating the reviewers' suggestions, test service personnel revised the specifications and prepared items which were tried out in several teacher training institutions in December 1939. Items described as "weak," "ambiguous," or "too easy or too difficult" were eliminated.[51] Following another review by practicing educators and final editing by test service personnel, the exams were ready for printing.

Based on the assumption that "admission to the profession in any capacity should be restricted to those above a certain minimum in intelligence, culture, and professional knowledge,"[52] a lengthy common examination was designed for all prospective teachers. Working time alone for it was eight hours. Total testing time for the combination of common and optional tests approached twelve hours, and two full days were required.

Despite the involvement of so many diverse groups in their planning and construction, the first National Teacher Examinations bore a striking resemblance to those related tests which preceded them and had precipitated their creation. They

[51] John Flanagan, "An Analysis of the Results from the First Annual Administration of the National Teacher Examinations," Journal of Experimental Education 9 (March 1941), p. 238.

[52] Wood, An Announcement of a Teacher Examination Service, p. 7.

certainly owed their titles, their comprehensive nature, and their exhaustive length to the exams developed by Ben Wood for the Pennsylvania study. Further additions and refinements had occurred before the inception of the National Teacher Examinations project, however. The Pennsylvania tests had been modified during the early 1930's by the Cooperative Test Service, and it was these new tests which were used "unofficially" to screen teaching candidates for the remainder of that decade. It was the anticipated discontinuation of this program which had prompted the superintendents to seek alternative testing arrangements and eventually to form the National Committee on Teacher Examinations. Consequently, it is not surprising that many sections of the teacher exams closely resembled the earlier tests.

Although a variety of "objective-type" items, including "true-false," "matching," and "fill-in-the-blank," had been included on the Pennsylvania and earlier Cooperative tests, all of the questions on the teacher examinations were "multiple choice" in nature. For all but a few items, candidates needed to select a single, correct or best, answer from five choices presented.

The Central Core of Knowledge for All Teachers: The Common Examination

The preparation of the common examination involved the development of fourteen distinct sections. The titles and contributing portions to the "common examination total score" were as follows:

1. Reasoning . . . . . 10 percent
2. English Comprehension . . . . . 10 percent
3. English Expression. . . . . 10 percent
4. Contemporary Affairs. . . . . 10 percent

General Culture:

5. Current Social Problems . . . . . 5 percent
6. History and Social Studies. . . . 5 percent

7. Literature. . . . . 5 percent

8. Science . . . . . 5 percent

9. Fine Arts . . . . . 5 percent

10. Mathematics . . . . . 5 percent

Professional Information:

11. Education and Social Policy . . . 7.5 percent

12. Child Development and Educational Psychology. . . . . 7.5 percent

13. Guidance, and Individual and Group Analysis. . . . . 7.5 percent

14. Elementary School Methods or Secondary School Methods. . 7.5 percent

Study of the candidates' practice booklet and of the tests themselves reveals an underlying triad. The common examinations were surely based upon those three "minimums" which Ben Wood assumed could be measured and had stressed that all teachers needed--basic intellectual and communication skills, cultural and contemporary background, and professional information. Wood's choices have shown remarkable longevity. Despite several intervening revisions, the current National Teacher Examinations also focus on these three areas.[53]

[53] Although the three separate tests which comprise the current core battery do not yield a combined "total score," they appear to sample content very similar to the original common exam and consist of the following three sections: (1) a test of communication skills--listening, reading, and both "multiple choice" and "essay" writing sections; (2) a test of general knowledge--literature and fine arts, mathematics, science, and social studies; and (3) a test of professional knowledge. [Educational Testing Service, NTE Programs: Core Battery Tests (Princeton: ETS, 1982).]

Assessment of "Intellectual Performance"

Those original tests included to measure the candidates' "general quality of intellectual performance"[54] were "Reasoning," "English Comprehension," and "English Expression." Designed to "sample quantitative non-verbal mental abilities,"[55] the Reasoning test was similar to the non-verbal components of other intelligence tests of the time, including those used earlier to evaluate students in teacher training institutions. In the "figure grouping" section, the candidate was required to indicate which of five geometric shapes did not possess a characteristic common to the other four. In the "number series" section, the test taker had to indicate which digits were missing from an arithmetic sequence. The "pattern analogies" task required candidates to select appropriate geometric figures to complete pictorial analogies.

The first of the two English tests, "English Comprehension," assessed reading ability. It combined forty "comprehension items based upon passages selected from books, magazines, [and] professional journals"[56] with a vocabulary test. Developed to measure skill in the correct use of the English language, the "English Expression" test included short sections on grammatical usage, punctuation, capitalization, spelling, "active vocabulary," sentence structure, and organization.

Assessment of General Cultural and Contemporary Background

Like those who had assessed normal school and college candidates earlier, Wood and the other NTE developers assumed that reasoning and verbal skills were key requirements for teaching. The knowledge they valued most, however, was that comprising the second major emphasis of the exams--knowledge of cultural and contemporary matters, of facts,

[54] National Committee on Teacher Examinations of the American Council on Education, Practice Booklet for Examinees (New York: Cooperative Test Service, 1940), p. 6.

[55] Ibid.

[56] Ibid., p. 7.

ideas, and concepts. The report of the Pennsylvania study, The Student and His Knowledge, had emphasized this kind of knowledge for it was highly valued by William Learned as well as by Ben Wood. Although "intelligence, an infectious personality . . . , serious and instinctive devotion . . . , depth and richness of experience, all contribute to produce a person from whom one can profitably hope to learn matters of importance . . . ," Learned had written in 1936, "a genuine teacher was never known to arise except from the soil of definite, abundant knowledge."[57]

Assessment in this component occupied half of the total working time, involved half of the total items, and contributed forty percent to the common examination total score. Included were the general culture test which concentrated upon "the broad perspective . . . fostered by a good general education"[58] and the test of contemporary affairs which assessed its current manifestations.

Clearly these exams were modeled almost directly after those of the same names developed previously for college students. The Cooperative General Culture test, first published in 1934 to sample "the traditional [content of] the liberal arts college curriculum,"[59] followed the precedent established in the Pennsylvania study and was composed of discrete subject matter exams.

The teacher test consisted of six sections, five of which were modeled after the previous cooperative versions--History and Social Studies, Literature, Fine Arts, Science, and Mathematics--and a sixth--Current Social Problems--which was added to the more general exam in 1941. Many of the teachers' items and virtually all of those in the Literature and Fine Arts sections required candidates to identify or define particular persons, objects, or terms by matching them with corresponding

[57] Learned, "Tested Achievement of Prospective Teachers," p. 31.

[58] National Committee, p. 8.

[59] F.S. Beers, "Cooperative General Culture Test," in The 1938 Mental Measurements Yearbook, Oscar K. Buros, ed. (New Brunswick: Rutgers University Press, 1938), p. 21.

characteristics or terms.[60] There were some questions which called for ability to draw inferences from facts, particularly in the History and Social Studies section. For example, in one of the sample questions, candidates needed to compare the problems of Britain "in the years following the World War" with "those of imperial Rome." Although the tests primarily sampled content taught in college, some of the items in each section but Mathematics dealt with knowledge of a contemporary origin or application. Some involved applying traditionally acquired facts to a contemporary situation.[61]

Dealing entirely with current situations was the Contemporary Affairs test. The original Cooperative version had been developed in 1934 at the University of Minnesota "to measure the extent to which college students keep abreast of significant current events."[62] "The content of the [original] test was selected on the basis of a survey of articles in twenty-five current magazines and journals . . . found to have the widest appeal on the junior college level."[63] Intended to be "a measure of the individual's acquaintance with current happenings,"[64] the teacher version emphasized very recent newsworthy content and was organized into three sections, one concentrating upon political news, the second on economic and social concerns, and the third on items from the fields of science, literature, and the arts. Unsurprisingly, about half of political items on the 1940 exam were military in content. Candidates were expected to know that "the present European war" had begun with an air raid on Poland and that the

[60] The following example was included in the candidates' practice booklet: "Chippendale is a name designating a style of---(1) architecture; (2) porcelain; (3) silverware; (4) furniture; (5) brocade."

[61] For example: "Destruction of the Panama Canal by bombers in time of war would be most serious to the United States in that it would . . . make it difficult for warships to get from the Atlantic to the Pacific."

[62] H.E. Hawkes, "Report on the Cooperative Test Service," The Educational Record 15 (July 1934), p. 366.

[63] Paul Limbert, "Cooperative Contemporary Affairs Test for College Students," in The 1938 Mental Measurements Yearbook, p. 67.

[64] National Committee, p. 10.

Graf Spee had been "scuttled by its crew."

Overall, the exams of general and contemporary culture emphasized the beliefs of the test makers that teachers need to possess a broad background of facts, ideas, and concepts from both traditional subject areas and contemporary social sciences. This comprehensive emphasis was especially indicated in revealing comments made by John Flanagan, associate project director, in his post-test analysis of the first examinations. Despite the test makers' disclaimers about potential influences on the training curricula, Flanagan indicated that although the reading comprehension test correlated highly with many of the other sections "it would be very dangerous to abandon the use of the other tests. If it were known in advance that the test contained nothing but Reading Comprehension, there would be grave danger that other factors in the training of candidates would be slighted."[65]

Assessment of Professional Information

Stressing that "only two hours of the total examination time of twelve" would be devoted to professional education subject matter, Wood's promotional announcement had de-emphasized the third test component.[66] Stating that "the committee [was] anxious to avoid any undesirable influence on the teacher training curriculum," he had reported that an emphasis would be placed on "functional knowledge" and "the common fundamentals."[67]

Determining the common fundamentals must have been a difficult task. Many teacher educators of the time believed that there was little professional agreement in this

[65] Flanagan, "An Analysis of the Results," p. 246.

[66] Wood, An Announcement of a Teacher Examination Service, p. 8.

[67] Ibid.

area.[68] Studies of existing programs found few common requirements.[69]

The exam included four newly created subtests. The first, designed to "measure the candidate's understanding of the implications of modern social trends for education and his familiarity with significant philosophical and historical developments relating to present day education,"[70] was called "Education and Social Policy." Primarily a test of the candidate's ability to recognize trends or "most commonly held" educational beliefs, it emphasized facts and concepts related to issues of social concern and social responsibility.

Candidates were asked to identify significant persons, agencies, and relationships. They were expected to know that John Dewey was "the chief contemporary exponent of the experimental method in philosophy" and that someone other than he was the U.S. Commissioner of Education. They needed to identify "the WPA," "the U.S. Office of Education," and the "4H." They were expected to know that the highest illiteracy rate was in the rural south and that constitutionally education was a function of the state. Candidates were asked to identify specific criticisms made by others of current and past practices. They needed to be familiar with the titles of a variety of progressive literature--Democracy and Education, The Child-Centered School, The Educational Frontier--but not necessarily with the application of principles espoused. Social policy was to be determined on the basis of factual information which was already established and recognized.

[68] For example, see Charles Judd, "The Curriculum of the Teachers College," in American Association of Teachers Colleges, Eighteenth Yearbook (Menasha, WI: Banta, 1926), pp. 75-82. Heilman (p. 73) believed this was the case even of fellow participants in the Teachers College Personnel Association's cooperative testing program.

[69] E.S. Evenden, "What Courses in Education Are Desirable in a Four-Year Curriculum in a State Teachers Colleges? What Should Be Their Scope?," in American Association of Teachers Colleges, Fifth Yearbook (Menasha, WI: Banta, 1926), p. 61.

[70] National Committee, p. 9.

The second test, "Child Development and Educational Psychology," covered that content common to most existing training programs and dealt "with the psychological background of classroom teaching and with the child's physical, emotional, and intellectual background."[71] Questions on social-emotional and personality development were generally related "functionally" to such matters as classroom control and "mental hygiene." Candidates were to be aware of adolescents' desire for "social approval" as "the most effective aid" in controlling them and to recognize that asking the pupils to "note something down" was a "desirable and effective method of recapturing the wandering attention of a class." Other questions inquired about children who were "shy and withdrawn," who used "bad language," and who demonstrated "persistent cruelty." Those items about intellectual development sampled such topics as intelligence testing, qualities of "precocious" and "exceptionally bright" children, and the adequacy of "traditional" school policies in dealing with high and low achieving pupils. About half of the questions dealt with what one of the items had defined as "the most important single topic in the field of educational psychology"--"human learning." Candidates were required to match the names of prominent psychologists to their major beliefs and to identify terms relating to past and currently popular views of learning.

As in the first section, most of the items were presented with a sense of unqualified certainty. The questions suggested the existence of a set of rules or prescriptions which candidates were expected to possess. It was implied that classroom decisions could be made on a standard or uniform basis if one knew the facts about "the best" or "the most effective" or "the most desirable" way to proceed.

The third section of the professional examination dealt with certainty of a different type--that gained from "the general principles of measurement . . . and the

[71] Ibid., p. 9.

interpretation of test results."[72] Although guidance and measurement were seen as closely entwined as the title "Guidance, and Individual and Group Analysis" implied, measurement alone was the focus of most items. In fact, only three dealt with principles of guidance. Another considered the superiority "in most respects" of children "of high intelligence," and the remaining fifty-six concerned aspects of "individual and group analysis." Teachers were expected to possess a great deal of technical knowledge about testing--knowledge which would allow them to construct their own exams and conduct their own testing programs as well to use the results of testing done by others. They were asked the meaning of "correlation coefficient," "percentile," "scatter diagram" and "norm" as well as the procedures for producing these from given data. They needed to interpret the results of testing--both generally and in "case study" type items in which the "best" academic and vocational decisions for fictitious students were to be selected on the basis of test data provided.

The final professional information test was the only portion of the common examination which was not completely "common" to all teachers. It was assumed that teachers preparing to teach at different levels would need to possess different knowledge of teaching principles. Thus, there were two versions of the final test--one of "elementary school methods" and one of "secondary school methods." Despite this distinction, however, the tests were quite similar. Both emphasized "present day developments in school management, classroom organization, and various aspects of the curriculum"[73] and included general questions about classroom management, planning and selecting materials, and evaluation. All candidates needed to know how to seat pupils with respect to "the principal source of daylight" and had to consider the influence a pupil's poor attendance record should have on his scholastic rating.

[72] Ibid.

[73] Ibid., p. 10.

Comments in articles published after the first administration indicated that, on the whole, the testers were quite pleased with the professional exam. It was felt that "the fact that the Reasoning test correlated more highly with [it] than any other of the common examinations except the Mathematics section" indicated "some confirmation" that it measured "reasoned understanding of school practices."[74] Also, it was believed that the wide range of candidates' scores indicated that there was "a genuine need among teacher training institutions for agreement on what constitutes an essential body of professional educational information and on methods for assuring that their students will have mastered at least this body of essentials before they become candidates for teaching positions."[75] That the multiple choice "one best answer" format might not be appropriate or that the test might reflect "questions of value" rather than "questions of fact" as has been suggested by critics of both the original and current exams[76] were not considered.

Conclusions

Although officially initiated with the March 1940 administration, impetus for the National Teacher Examinations project had been developing for several decades prior to that date. The testing program evolved from a number of related factors including the past use of examinations for teacher certification, concern about standards and growing reliance upon testing in many phases of education, and an existing surplus of teachers which was believed to vary considerably in ability and training. The availability of

74 William Crissy, "The National Teacher Examinations," Phi Delta Kappan 24 (May 1942), p. 355.

75 Frederick Davis, "The Measurement of Professional Information among Candidates for Teaching Positions," Educational Administration and Supervision 27 (February 1941), p. 106.

76 See, for example, John Pilley, "The National Teacher Examination Service, " School Review 49 (March 1941), pp. 177-186 and Marleen Pugach and James Raths, "Testing Teachers: Analysis and Recommendations," Journal of Teacher Education 34 (January-February 1983), pp. 37-43.

private foundation funds with which to subsidize the original project and the existence of prepared examinations which had been used previously to screen teacher candidates were additional contributing factors.

Reflected in those first examinations are assumptions about teacher competence, assessment, and knowledge which were held by test developers of that period and which have influenced teacher testing since that time. Clearly, Ben Wood and his collaborators sought to reform both the schools and the larger society. The means to this end, they believed, was the improvement and upgrading of the teaching force. Practicing teachers were assumed to possess "mediocre minds," to be unintelligent, un- or under-educated in the liberal arts, and poorly or incorrectly trained in pedagogy. Similarly, those responsible for existing teacher education programs were seen as ineffective, misguided, or incompetent. Some critics of that time, like Goodwin Watson, called for raising teacher salaries to enhance educators' public esteem and to attract candidates of "higher quality." However, many reformers, then like now, saw teacher testing as a necessary and laudable measure. Like present-day school administrators and public officials who seek to control who teaches in schools, Ben Wood and his supporters saw themselves as strengthening both teaching and teacher education through the creation of a common, although admittedly minimal, standard.

Some variations in the foci and perimeters of that standard have occurred in the past five decades. Certainly, the State of Arkansas's controversial mandated program of minimal competency testing and Albert Shanker's proposal for a "rigorous" national professional examination reflect current variations. However, the assumptions that significant school and societal reforms will be achieved through "the improvement" of a mediocre teaching force and that this will be accomplished by the administration of paper and pencil tests focusing upon "minimal" basic skills, "elite" liberal arts knowledge, and/or "established" professional or pedagogical information have persisted.